MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Non-monotonic power in Bayesian dynamic borrowing: insights and practical remedies

Gianmarco Caruso* and Pavel Mozgunov

PSI 2025 conference, Wembley Stadium, London, UK
June 11, 2025

*gianmarco.caruso@mrc-bsu.cam.ac.uk

Motivation: non-inferiority oncology trial

RCT to compare two different therapies for treating a cancer that originates in the appendix:

- High-dose regimen ("Dutch protocol", most used)
- Low-dose regimen (cheaper and *safer*, used in the UK centre)

Can we conclude that the low dose is **not less effective** than high dose?

Endpoint: **2-year disease-free survival (DFS)**

$$Y_j \sim Bernoulli(p_j), \qquad j = \{High, Low\}$$

with Bayesian priors on survival rates p_j's based on historical data

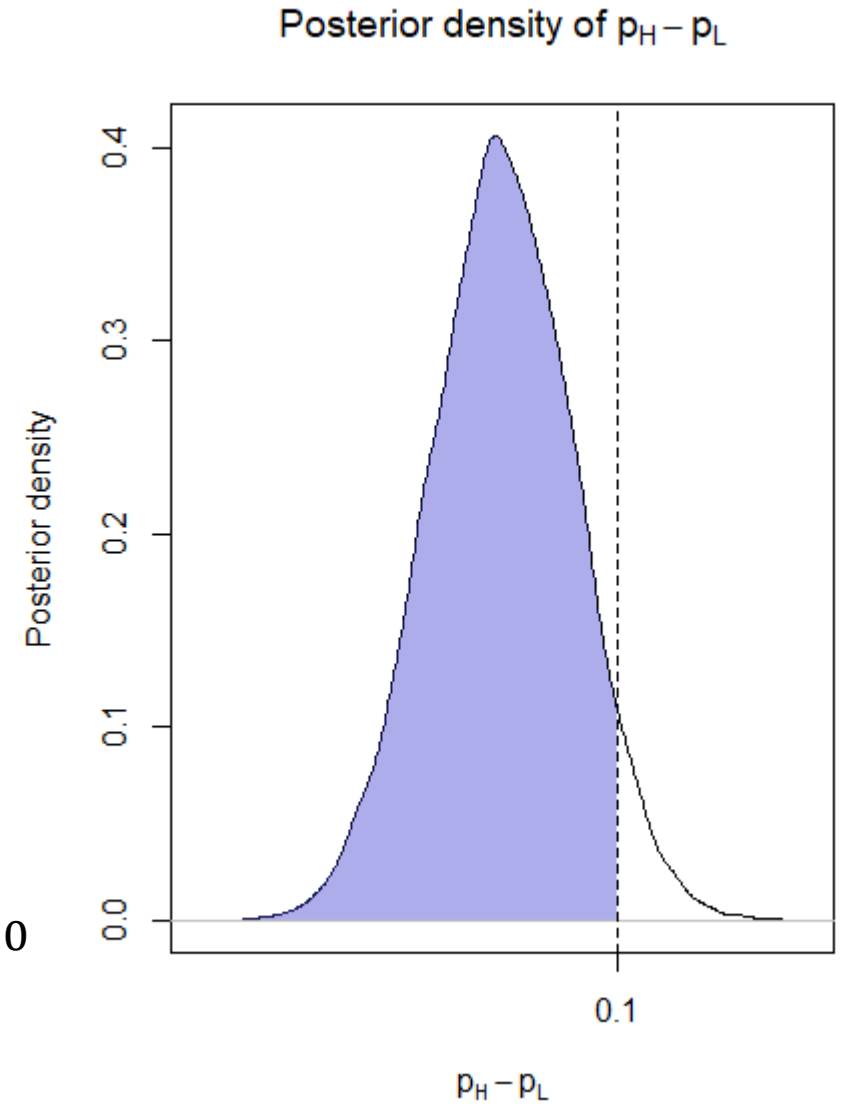
Bayesian decision-making

Hypotheses of non-inferiority trial:

- H_0 : Difference in survival rates ($p_H - p_L$) is **larger or equal to** 0.1 (**low-dose is inferior)**
- H_1 : Difference in survival rates ($p_H - p_L$) is **smaller** than 0.1 (**low-dose is not inferior)**

We claim non-inferiority of low-dose regimen if $P(p_H - p_L < 0.1 | data) > \xi$

N.B. ξ calibrated to control type I error under specific H_0



MRC Biostatistics Unit

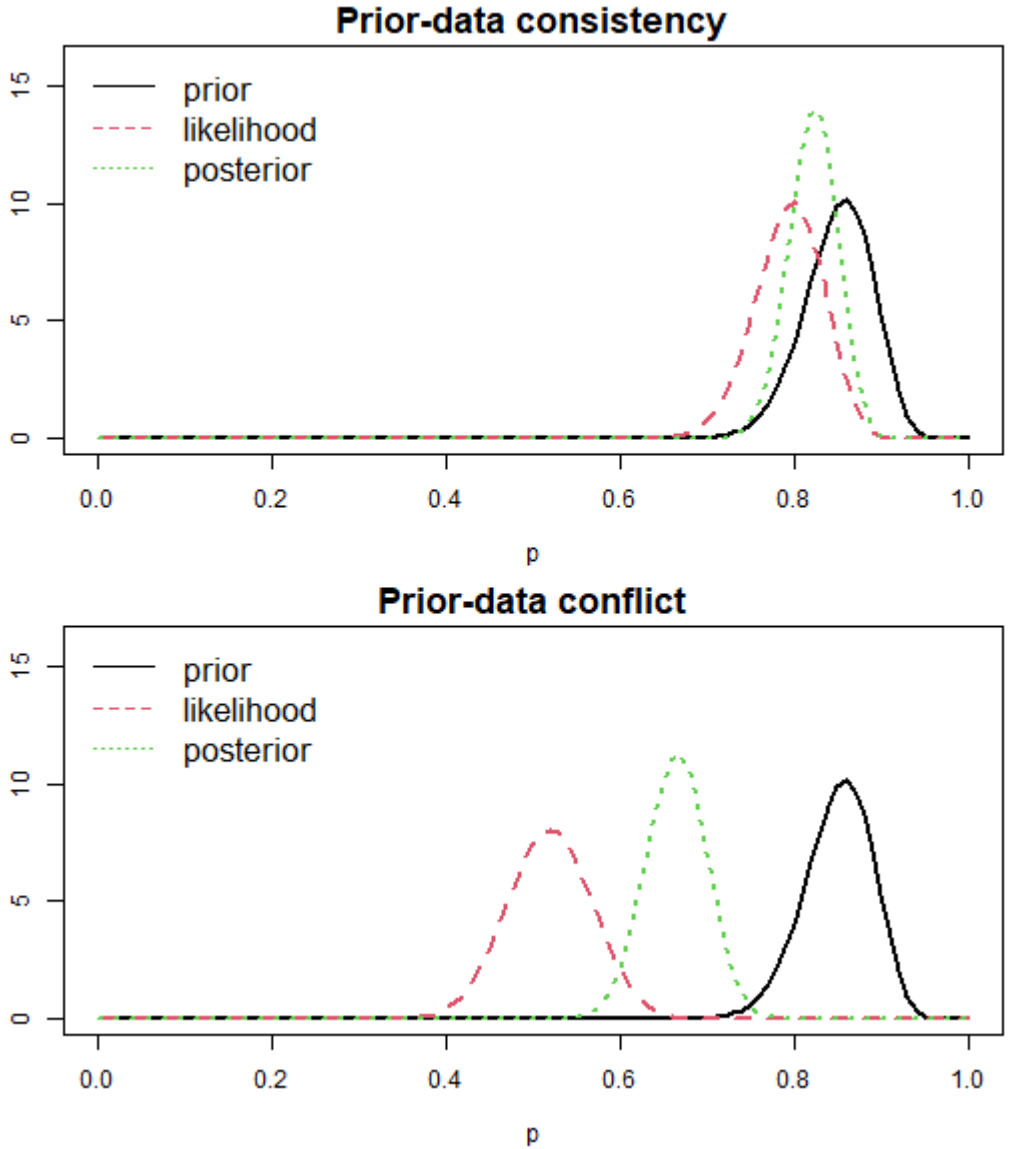
UNIVERSITY OF CAMBRIDGE

Prior elicitation from historical data

Prior distributions based on historical survival rates and degree of relevance of past studies (conjugate Beta, MAP, etc.)

Benefit: If new and past study data are ***consistent,*** we gain **efficiency**
(=smaller sample size to achieve same power)

Risk: In case of **conflict**, a too informative prior can **increase the chance of making wrong conclusions (!)**



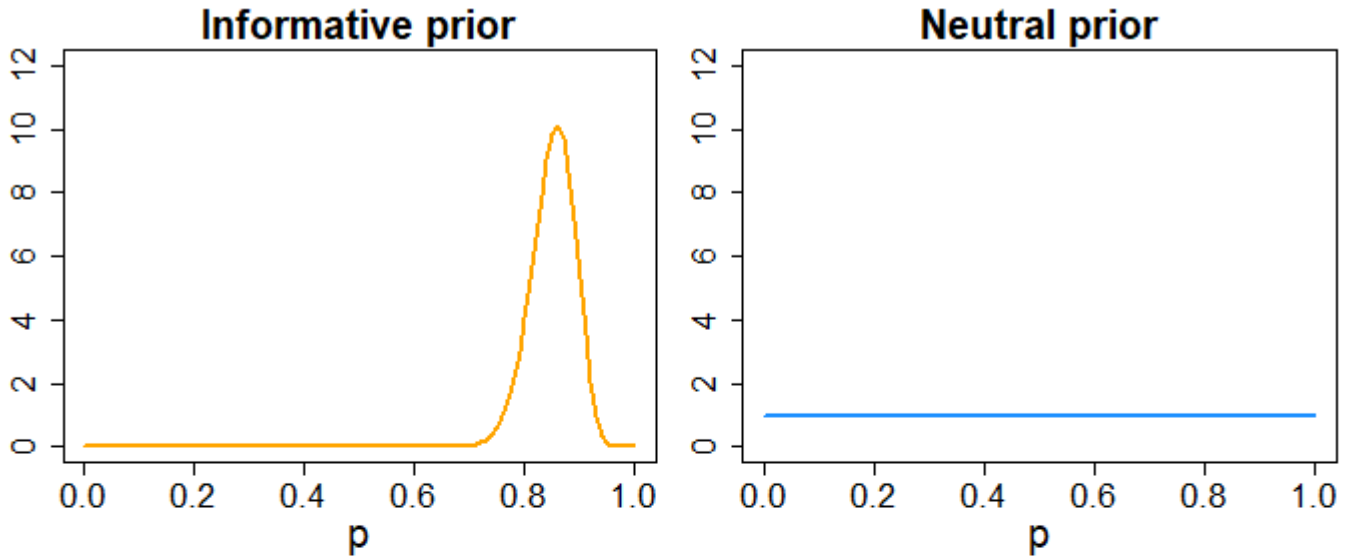
MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Robustification: get prepared for the unexpected

Solution[1]**:** *mix* informative prior (from past data) with a neutral one

$$\pi_R(p) \;\xrightarrow{w_r}\; \pi_1(p) \qquad \pi_R(p) \;\xrightarrow{1-w_r}\; \pi_0(p)$$



$\boldsymbol{w_r} \in [\mathbf{0}, \mathbf{1}]$ reflects how much we rely on past data

The smaller $\boldsymbol{w_r}$, the less rely on the past
…but the potential of Bayesian approach is reduced **(!)**

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

1. Schmidli et al. (2014)

Self-Adapting Mixture (SAM) priors

SAM prior[2] dynamically adapts the level of borrowing from historical prior based on the *degree of consistence* between historical ($\mathrm{D_h}$) and new data (D).

$$p \sim \tilde{w} \cdot \pi_1(\theta) + (1 - \tilde{w}) \cdot \pi_0(\theta)$$

informative prior based on $\mathrm{D_h}$ and with mean θ_h

Uniform[0,1]

If $H_0: \theta = \theta_h$ (D and $\mathrm{D_h}$ **consistent**) and $H_1: \theta = \theta_h \pm 0.05$ (D and $\mathrm{D_h}$ in **conflict**), then

$$\tilde{w} \propto \frac{L_\theta(\mathrm{D}|H_0)}{L_\theta(\mathrm{D}|H_1)} = \frac{L_\theta(\mathrm{D}|\theta = \theta_h)}{\max\{L_\theta(\mathrm{D}|\theta = \theta_h \pm 0.05)\}}$$

The larger the conflict between D and $\mathrm{D_h}$, the less the borrowing from $\mathrm{D_h}$

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

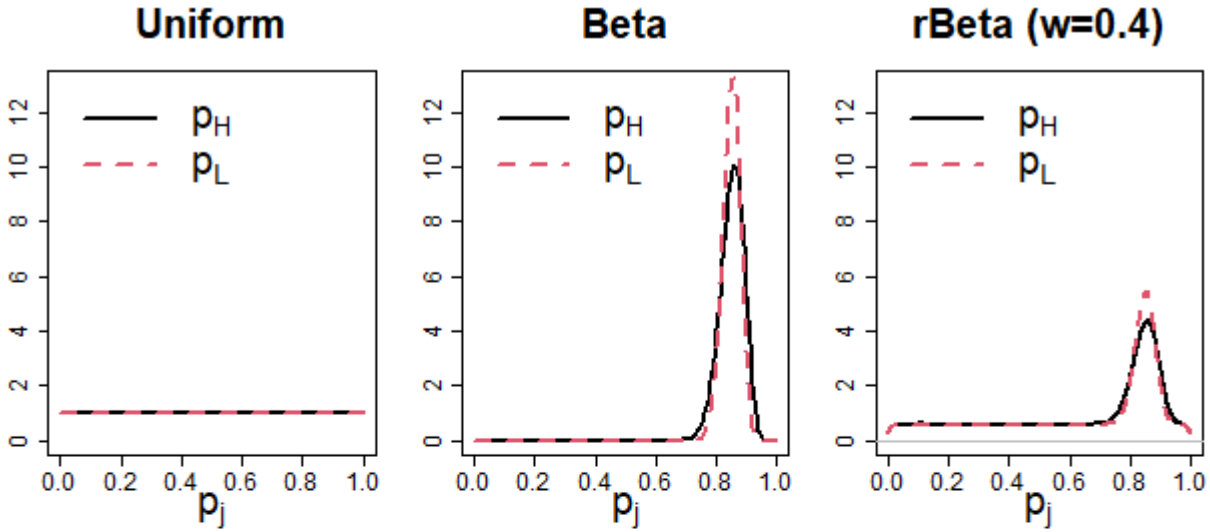
2. Yang et al. (2023)

Prior specifications (example)

- **Uniform:** $p_H, p_L \sim Unif\,[0,1]$
- **Beta:** $p_H \sim Beta\left(mean = \hat{p}_j^H\,, sd = 0.04\right)$

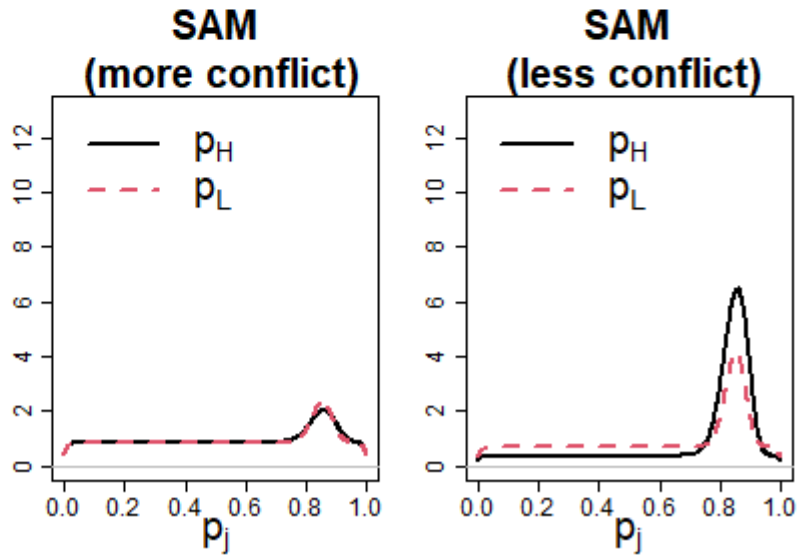
  $p_H \sim Beta\left(mean = \hat{p}_j^H\,, sd = 0.03\right)$

  with $\hat{\boldsymbol{p}}_{\boldsymbol{j}}^{\boldsymbol{H}} = \mathbf{0.85}$ **historical survival rate**



- **rBeta:** $p_j \sim w \cdot Beta_j + (1-w) \cdot Unif,\ w = 0.4,\ j = \{H, L\}$
- **SAM:** same as rBeta but with **dynamic** weight w_j determined by the level of prior-data conflict

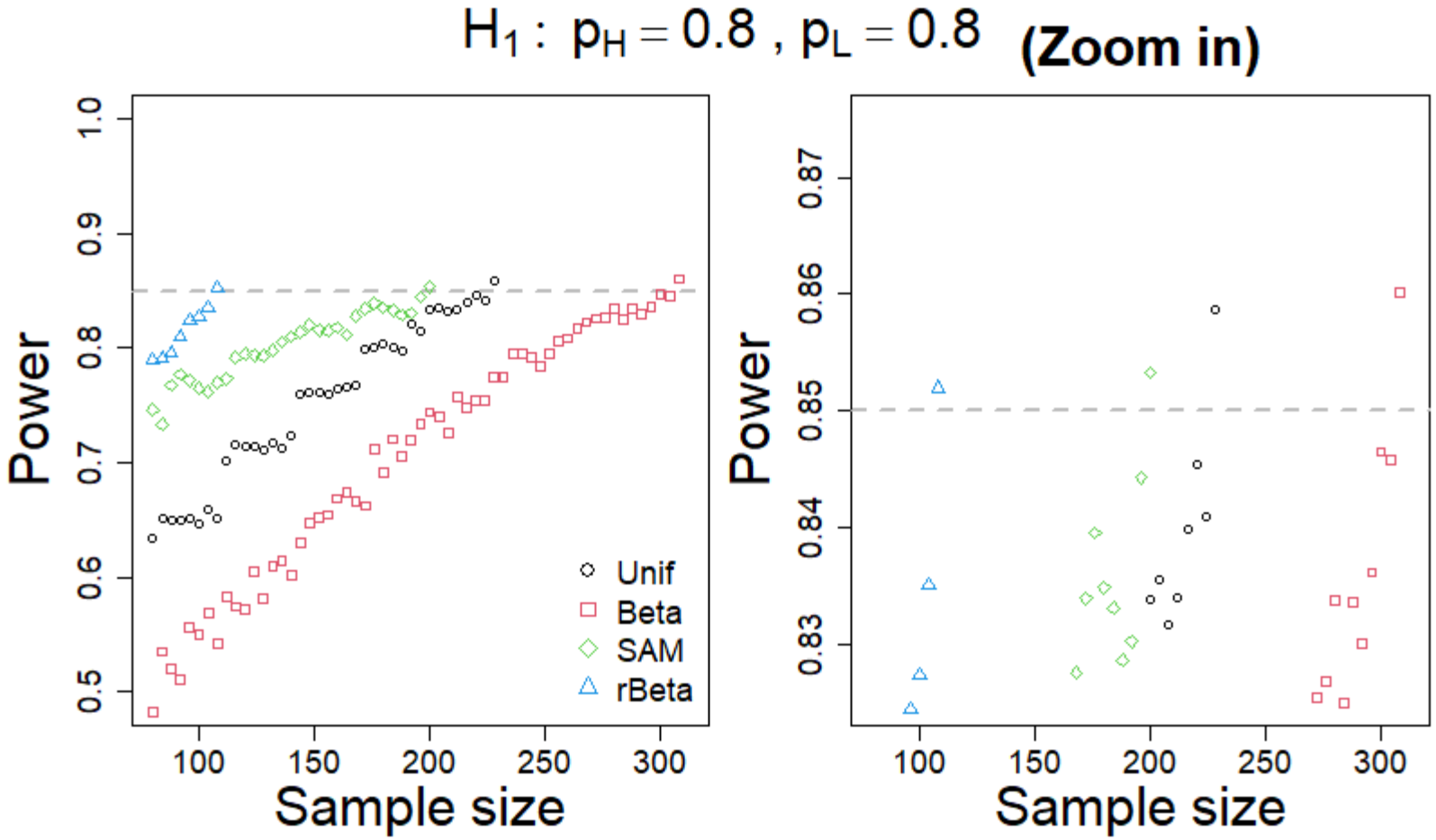
  (N.B. the larger $|\hat{p}_j - \hat{p}_j^H|$, the larger the conflict)



MRC Biostatistics Unit
UNIVERSITY OF CAMBRIDGE

Sample size determination

- Forward search on grid of sample sizes $(n_C = n_L = n/2)$
- Stop when target power (85%) is achieved
- ξ calibrated to control $\alpha = 20\%$ under $H_0: p_H = 0.75, p_L = 0.65$



MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

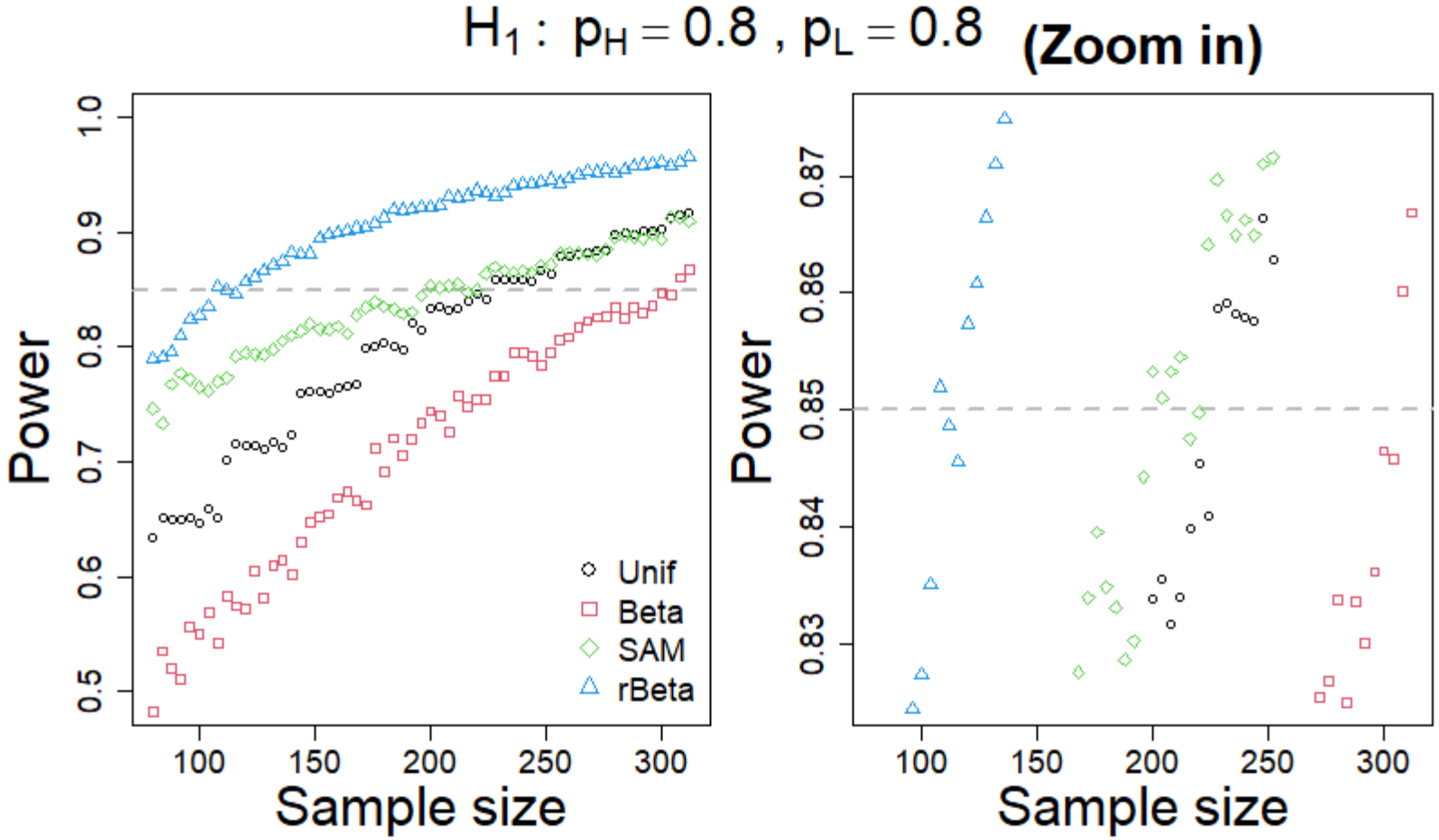
Results based on 50k posterior samples for each of the 200k replicas

Sample size determination

Caution:

Estimated power may decrease with sample size (simulation error?)

Counterintuive conclusion if we only rely on the power for a specific sample size



MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

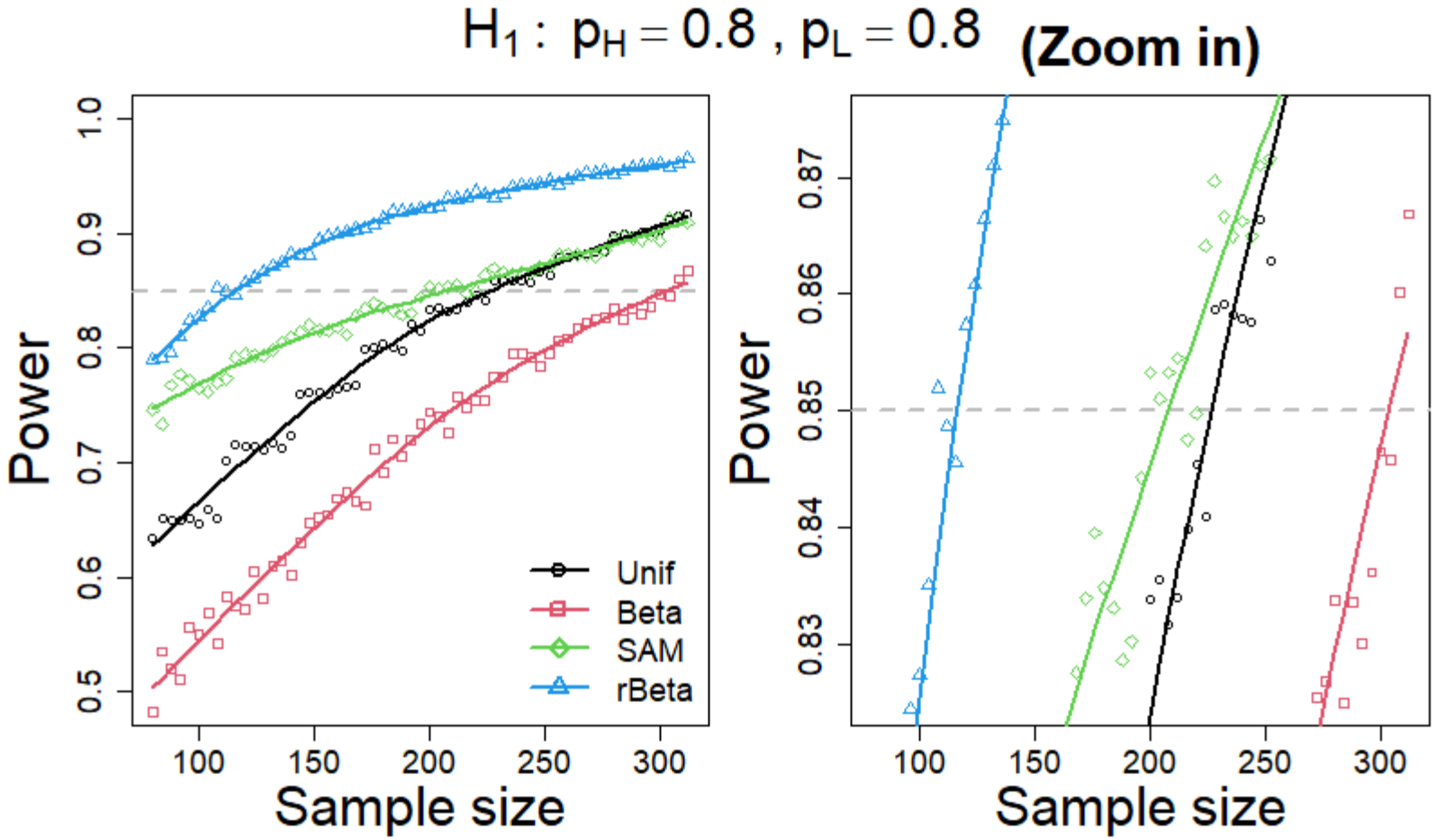
Results based on 50k posterior samples for each of the 200k replicas

Sample size determination

Solution:

Non-parametric regression to fit smooth power curve

More coherent interpretation of power vs sample size relationship



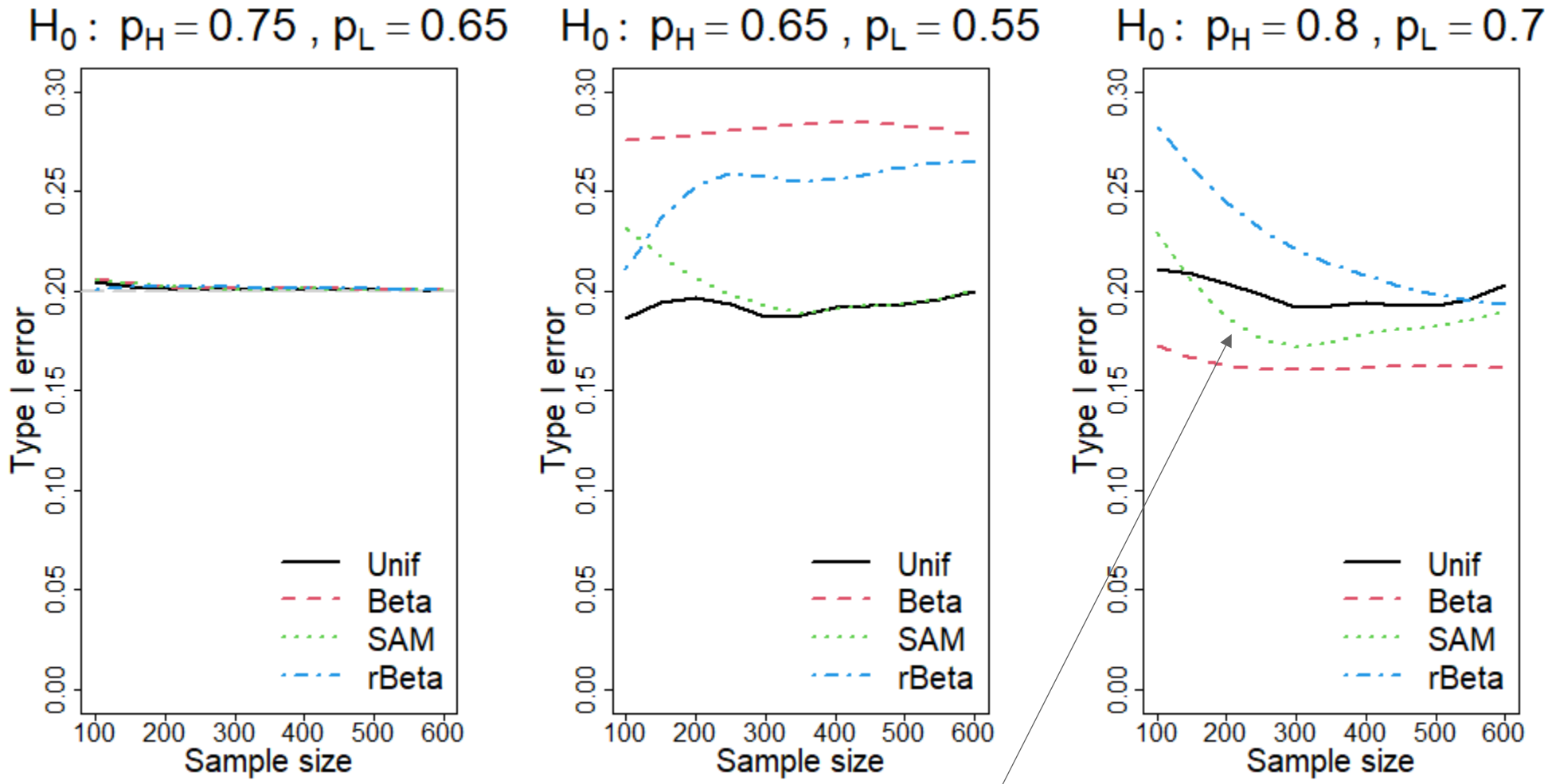
MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Results based on 50k posterior samples for each of the 200k replicas

What if the overall trend itself is non-monotonic?

Type-I error rates

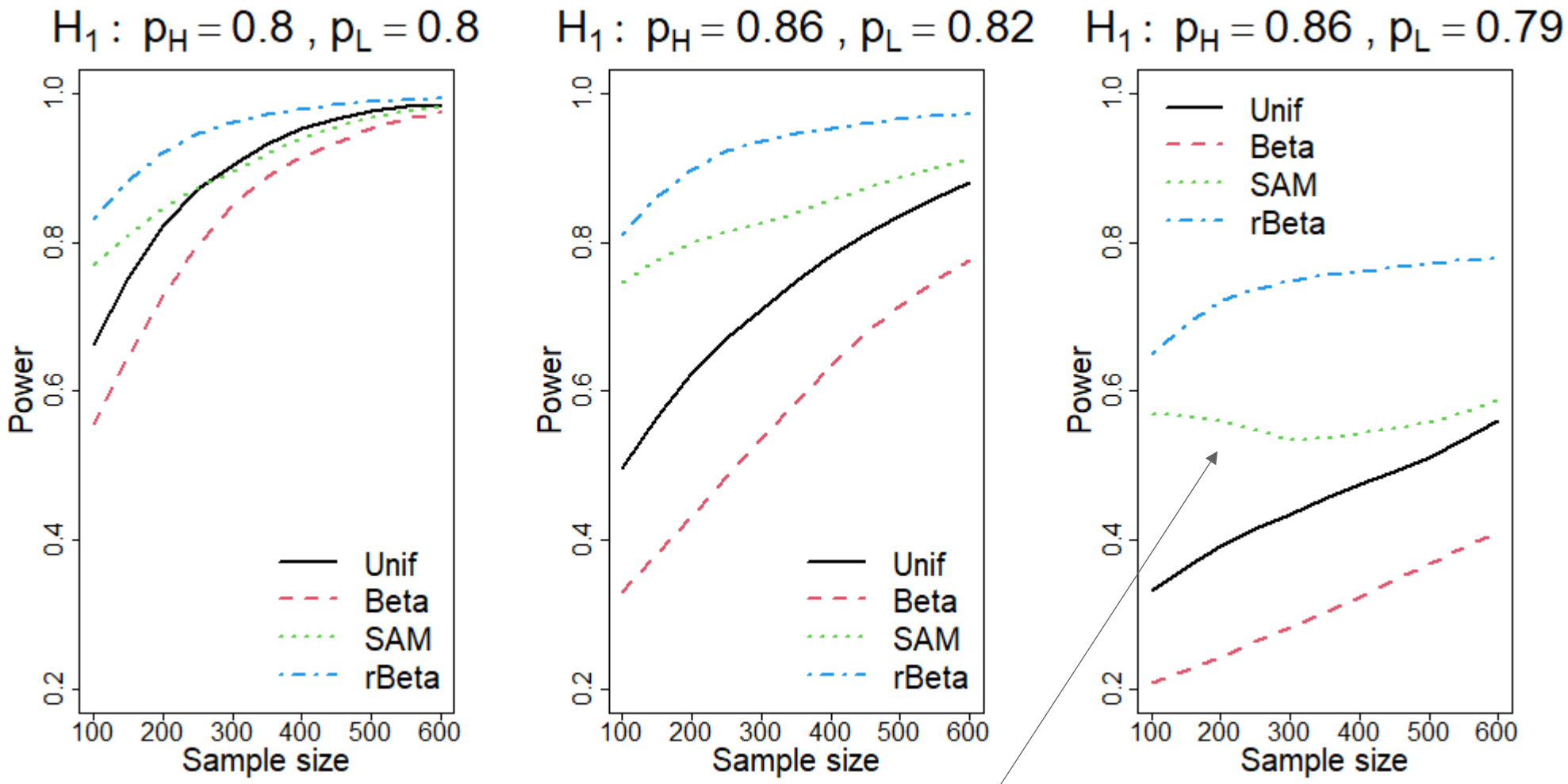


Non-monotonic trend of type-I error

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Power



Non-monotonic trend of power

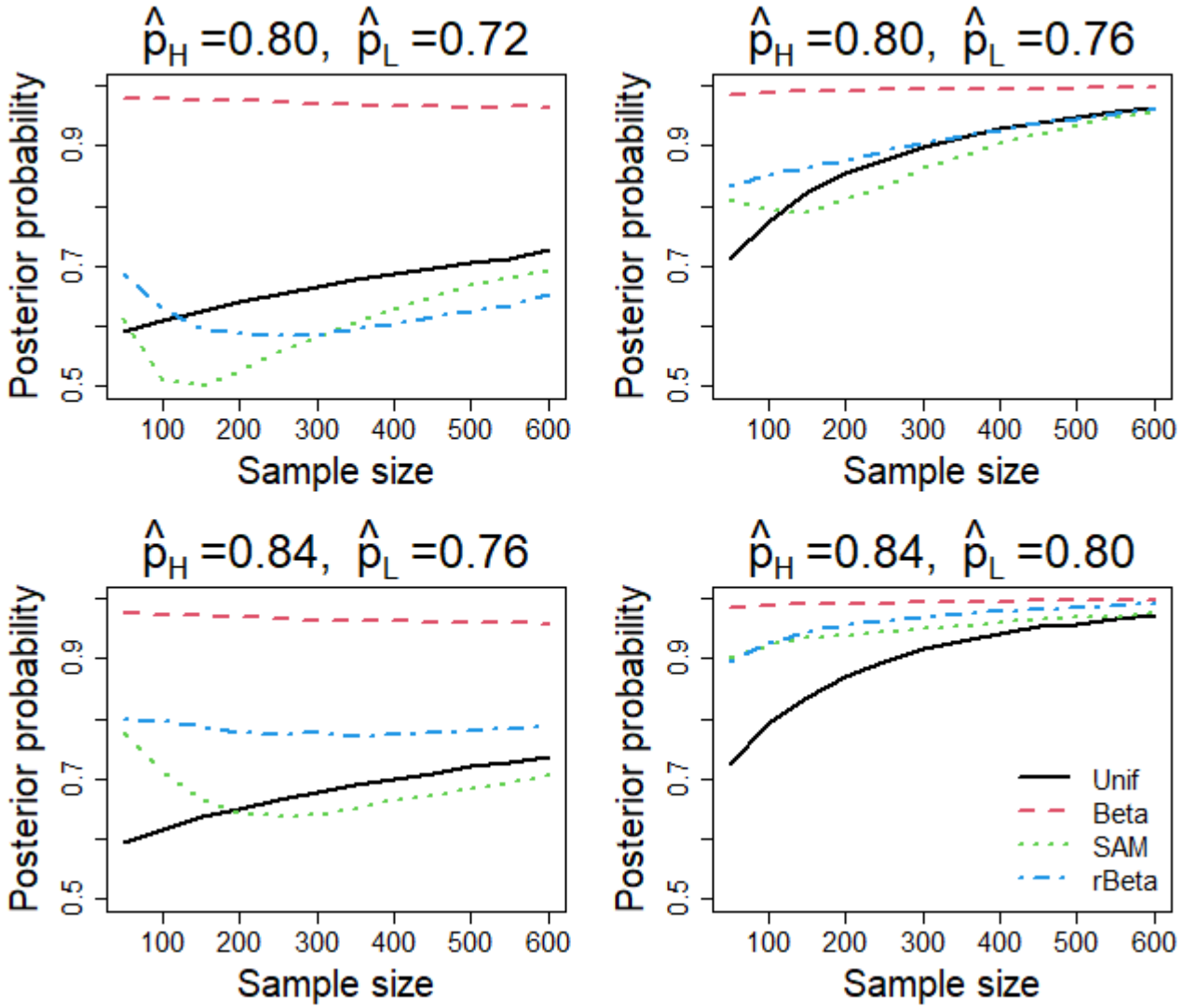
MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Closer look at posterior probability of H_0 being false

For fixed observed response rates $(\hat{p}_H, \hat{p}_L)$, posterior probability that H_0 is false based on SAM prior shows **evident non-monotonic behaviour** with sample size

Similar result obtained by Kopp-Schneider et al. (2020) on single-arm trial, with posterior probability decreasing with number of responses



MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Final remarks

Dynamic borrowing of historical information allows for **prompt reaction to prior-data conflict**, by discounting the amount of information borrowed

…but careful use is needed:

- Bayesian **principle of independence** between prior and data is **violated**
- In some scenarios, the conflict between informative prior component and data may lead to **power curves that are non-monotonic with sample size**
- Sensitivity analysis for different sample sizes may be needed, above all **if more patients then expected are recruited**

Main references

Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D., & Neuenschwander, B. (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics*, *70*(4).

Yang, P., Zhao, Y., Nie, L., Vallejo, J. & Yuan, Y. (2023). SAM: Self-adapting mixture prior to dynamically borrow information from historical data in clinical trials. *Biometrics*, 79(4).

Kopp-Schneider A, Calderazzo S, Wiesenfarth M. (2020). Power gains by using external information in clinical trials are typically not possible when requiring strict type I error control. *Biometrical Journal*, 62(2).

MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Thank you for your attention!